Optimization Algorithms
Exercise 9

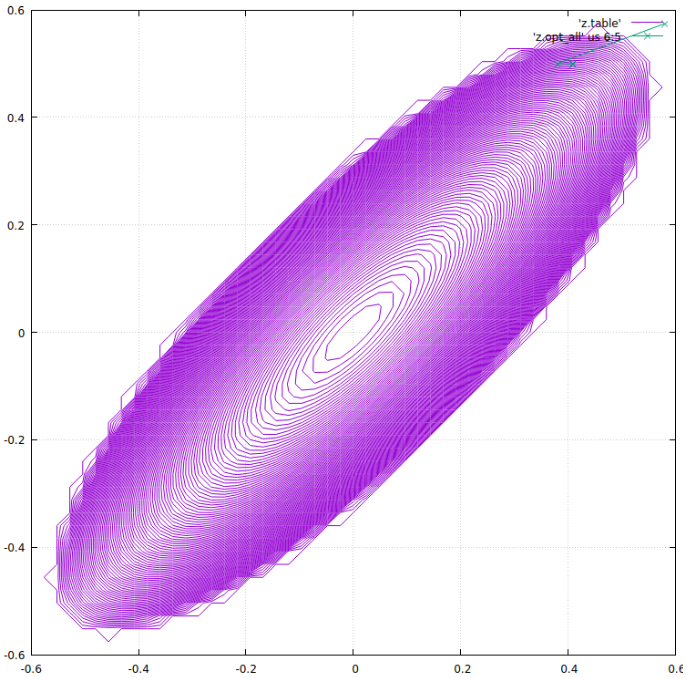
Marc Toussaint
Learning & Intelligent Systems Lab, TU Berlin
Marchstr. 23, 10587 Berlin, Germany

Winter 2020/21

1 Bound Constrains

The demo in the exercise considered the problem: $x \in \mathbb{R}^2$

$$\min_x \frac{1}{2} x^\top A x \quad \text{s.t.} \quad x_2 \geq \frac{1}{2} \ , \quad \text{with } A = \begin{pmatrix} 200 & -160 \\ -160 & 200 \end{pmatrix}$$



a) Analytically compute the optimum for this problem. (For arbitrary A, the specific numbers are not important.)

b) Assume we are at location $x = (0, 1)$. In which direction does the gradient $-\nabla f$ point, and in which direction does the Newton step $-\nabla^2 f^{-1} \nabla f$ point? (Compute them, and illustrate them in a picture.)

c) Assume we initialize our bound constrained Newton method (slide 9/18 of the lecture) at $x = (0, 1)$, how many Newton iterations (where each iteration does line search in the determined direction δ), will it need until convergence. Illustrate roughly, where each step moves to.

d) Let us define $r(x_1) = f(x_1, x_2 = \frac{1}{2})$, which is the cost function on the hyperplane only. Given any point x_1 on the hyperplane, what is the Newton step within the hyperplane w.r.t. x_1? Is this the same as the (clipped) Newton step for $f(x_1, x_2)$ when deleting the off-diagonal terms from A (as our method does)? Prove.

2 Restarts

Consider the Rastrigin function with $f(x) = ad + \sum_{i=1}^{d}(x_i^2 - a\cos(2\pi x_i))$ with $x \in \mathbb{R}^d, a = 6$ and $d = 2$ or $d = 10$. The global minimum of this function is at $x = 0$. It is very similar to the function you considered already in sheet e03, exercise 3.

a) Plot the function for the 2D case, $x \in \mathbb{R}^2$.

b) Implement a simple random restart mechanism for a Newton solver.

c) Plot the best function value obtained so far over the number of restarts.

d) Can you come up with a better sampling strategy than uniform sampling?